

PROPOSAL OF FINAL PROJECT

year 2021–2022

Research group or lab: Evolutionary Genomics and Bioinformatics

Institution: Universitat de Barcelona

Supervisor(s): Julio Rozas / Alejandro Sánchez-Gracia

Contact: jrozas@ub.edu

Title of the project: Comparative and evolutionary analysis of multi-gene families in spider genomes

Tasks: The student will participate in the identification, annotation and evolutionary analysis of different gene families in complete genomes. For that, he/she will use high quality genome sequences, bioinformatics tools (software and scripts to manipulate and visualize sequences and genomic annotations, to identify gene family copies), and will analyse the data using phylogenetic and evolutionary genetic methods. Many of these analyses will be carried out in our high performance computer cluster.

Physical work address: Facultat de Biologia; Departament de Genètica, Microbiologia i Estadística

Training requirements: Notions and knowledge about comparative genomics and phylogenetics, and/or on NGS data handling, assembly and analysis. Experience with Linux operating systems, and some programming languages commonly used in bioinformatics (Perl, Python, R) are desirable.

Description of the project:

Understanding the origin and evolution of gene families is a central question in Evolutionary Biology. Despite that current high-throughput sequencing (HTS) technologies are accessible for many labs, the accurate identification and annotation of gene family members in genome assemblies of non-model organisms is a major challenge. This scenario will change in the near future with the irruption of the so called “Third-generation sequencing technologies”. Currently, our research group generates genomic and transcriptomic data of spiders (Chelicerata), including DNA sequence data using long reads and Hi-C, Chicago and Omni-C libraries.

The objectives of this TFM is to perform a comparative genomic and molecular evolution study of some multigene families in chelicerates and, by extension, in arthropods. Among others, we are currently studying gene families encoding i) venoms and toxins, ii) tRNAs, iii) involved in the chemosensory system (olfactory and gustatory). These systems have very relevant biological

each characteristics with many applications and interest beyond evolutionary biology. For that, we are using comparative genomics and transcriptomics approaches, under the theoretical framework of molecular evolutionary genetics to identify the genomic regions and gene functions driving diversification. We applied powerful bioinformatics tools to detect changes in coding and non-coding regions, in gene copy number, and in gene expression levels associated with speciation processes.

The student will participate in the assembly, annotation and analysis of the spiders *Dysdera silvatica* and *D. catalonica* genomes. At this time, we have both genomic (genome assembly using short, long reads, and Hi-C data), and transcriptomic (RNAseq data). The student will learn and use specific software for genome/transcriptome assembly, for gene annotation, and for conducting comparative genomics and phylogenomics analyses.

Comments: